Entanglement-Based Quantum Key Distribution

Daniel Swift

Project Supervisor: Prof. Paul Busch

Abstract

At the heart of cryptography is the problem of establishing secure communications between two parties over public channels. Over the past few decades, increases in computing power has resulted in cryptosystems, once assumed to be secure, being broken. The advent of quantum computing presents new challenges; another increase in computing power and the ability to implement the highly efficient Shor's algorithm. This report presents an entanglement-based quantum key distribution (QKD) protocol, known as Ekert 91 which can be used to implement the provably secure one time pad. The report begins by reviewing where current public key methods such as RSA are failing and introduces QKD as a solution. This then leads into a mathematical introduction detailing key concepts such as quantum state discrimination, quantum entanglement and the positive partial transpose criterion. Next, we consider the famous EPR paradox and derive the generalised Bell inequality known as the CHSH inequality, which is then used to test for eavesdropping in the Ekert 91 protocol. After this, an example of an eavesdropping attack is presented where Eve prepares two particles with an ancilla and imitates environmental noise to avoid detection. To minimise errors associated with state discrimination an optimal strategy is presented using the Helstrom bound. Finally, it is shown that secret keys can be established despite an eavesdropping attack taking place, this is achieved by using purification-based quantum privacy amplification.

Acknowledgements

I would like to thank my project supervisor, Professor Paul Busch for sharing his expertise and providing support throughout the project. I would also like to express my gratitude to my family and friends for supporting my pursuit of higher education. Thanks are also due to the so-called "*Mathsters*" group for their encouragement and late night discussions, and especially to Alex Batteson and Nora Gilbertson for their help in proof reading this report.

Contents

List of Figures

List of Tables

1 Introduction

At the heart of cryptography is the problem of establishing secure communications between two parties, commonly referred to as Alice and Bob. Alice and Bob require a method of encrypting their communications in such a way that a third party or eavesdropper, commonly referred to as Eve, cannot decrypt or gain any useful information [1]. For as long as there have been code-makers, there have also been code-breakers. Therefore, the cryptographic methods employed by Alice and Bob need to be able to withstand attacks from malicious third parties.

The majority of cryptographic methods presently used to secure communications over public channels rely on computationally hard problems. For example RSA, named after the surnames of its creators, Rivest, Shamir and Adleman, depends on the difficulty of factoring large prime numbers and Diffie-Hellman-Merkle (DHM) depends on the discrete logarithm problem. Cryptographic methods that rely on computationally hard problems fall victim to advances in algorithms designed to solve these problem and even more so to advances in computing power [1]. In fact, in 1977 Rivest stated that based on computers and factoring methods of the time it would take 40 quadrillion years to crack an RSA-125 key. RSA-129 was cracked just 17 years after this claim was made due to advances in computing power; the current recommended key size for the implementation of RSA is 2048 bits which is expected to be secure until 2030 [2].

With the recent advent of quantum-based computing we will see another leap in computing performance. With quantum computing there is a class of quantum algorithms which are likely to be more efficient than their classical analogues. Most notably there is Shor's algorithm [3] which can be used to decompose an integer N into its prime factors. This is the exact problem that RSA is built upon. Shor's algorithm is much faster than the general number field sieve, the previous most efficient method. In terms of cryptographic methods, we need to use computationally harder problems or consider completely different methods for securing communications.

Quantum key distribution (QKD) is one solution to the problem outlined above. QKD uses features of quantum mechanics such as the superposition principle, Heisenberg's uncertainty principle and entanglement to secure lines of communication over public channels by generating a shared secret key between two parties. The first instance of associating a notion of secrecy to quantum theory was established by Wiesner in the 1970s [4] and published in 1983. Wiesner was considering the idea of quantum money which would be impossible to counterfeit. The first QKD protocol was established by Bennett and Brassard in 1984 [5] and is based on disturbance of non-orthogonal states; the protocol is known as BB84. In 1991, Artur Ekert proposed an entanglement-based QKD protocol [6] which will be the focus of this report. Unlike non-quantum key distribution methods, variants of the BB84 and Ekert 91 are provably secure, and not dependent on computationally hard problems.

In this report we will swiftly review RSA, a public key cryptosystem in use today and explore the premise of its computational security motivating the use for alternate methods of key distribution. We will then explore the entanglement-based quantum key distribution protocol known as Ekert 91 and observe how the non-locality of quantum theory provides methods to test for eavesdropping. To achieve this we will review the Einstein-Podolsky-Rosen paradox [7] and observe quantum violations of a generalised Bell inequality known as the Clauser-Horne-Shimony-Holt (CHSH) inequality [8, 9].

We then explore quantum privacy amplification (QPA) which is a method for reducing the amount of information Eve can learn about Alice and Bob's shared key. We will investigate how purification based QPA [10] can be applied to generate secure keys for non-optimal violations of the CHSH inequality and more impressively non-violations.

1.1 Significance of this work

QKD is a rich subject area that encompasses aspects of cryptography, information theory and quantum mechanics. This can make the subject inaccessible to those unfamiliar with any of these areas. The basic notion of protocols such as Bennett and Brassard's BB84 are generally well covered at undergraduate and M-Level, however, information on entanglement-based quantum key distribution is not. Thus, this project aims to provide a clear and concise introduction to entanglement-based QKD. Given the sheer breadth and depth of QKD as a field it is difficult to cover everything in detail, especially some of the subtleties and nuances associated with protocol security. In most literature, eavesdropping and quantum privacy amplification (QPA) tend to be non-mathematical or inaccessible to undergraduate readers. This project has sought to explain an eavesdropping attack and application of QPA in a mathematically approachable way that an undergraduate reader can appreciate. In summary, this project aims to detail key components of entanglement-based QKD and link them to the wider research as a basis for readers to gain the necessary knowledge to approach further research.

1.2 Outline

Throughout this report we will assume that the reader has knowledge of linear algebra and a basic familiarity with number theory and quantum mechanics. The focus of this report is on how entanglement-based quantum key distribution can be used to establish secure communication over insecure channels.

In section 2 we begin by briefly reviewing historical developments in the field of cryptography. We then introduce public key cryptography and focus on the the problem of establishing secret keys over public channels. We will also review the computationally-hard problem of factoring large integers, the problem that RSA relies on, and review the advances in solving this problems. Lastly, we will discuss how attacks using quantum factoring algorithms such as Shor's algorithm will further exacerbate this issue.

In section 3 the mathematical background to understand the following sections will be introduced. This will include Hilbert spaces, operators, and quantum mechanics. We will introduce quantum entanglement and give examples of how to check whether a state is entangled or separable.

In section 4 we will discuss the historical context and implications of entanglement and review its rich history including the EPR Paradox and Bell's Theorem which states that no local hidden variable theory can reproduce all the predictions of quantum mechanics. Lastly we will derive a generalised Bell inequality known as the CHSH inequality and explain the implications of quantum violations of the inequality.

In section 5 a general outline of a QKD protocol will be provided and important results such as the No-Cloning theorem and how information gain implies disturbance will be discussed. After this we will focus on an entanglement-based protocol known as Ekert 91 and the advantages of using entanglement-based QKD. We will show that CHSH violations can be used to test for eavesdropping and an eavesdropping attack will be demonstrated in which Eve entangles two particles with an ancillary particle. We then discuss entanglement purification and introduce purification-based QPA to show that secure communications can be established under an eavesdropping attack.

2 Past and Present Cryptography: What is secure?

The Caesar cipher was used over 2000 years ago and is one of the first documented cryptographic methods. The Caeser cipher indexes letters of the English alphabet from 0 to 25, where 0 and 25 correspond to a and z respectively. Plain text letters are encrypted by moving them 3 places to the left in the alphabet. 2000 years later, we can see that shift ciphers are completely insecure. One merely needs to try 26 shifts to decrypt the message by brute force.

The Vigenère cipher is a notable improvement on the Caesar cipher. The cipher encrypts plain text by repeating a key until all of the plain text has been encrypted. The Vigenère cipher is a poly-alphabetic substitution cipher meaning that different plain text letters can be encrypted to the same cipher text. This is demonstrated in table 2 where letters of the plain text message can be seen to map to the same letter in the cipher text, for example $h \mapsto V$ and $e \mapsto V$. Compared to shift ciphers where there is a 1-1 mapping between plain text and cipher text, poly-alphabetic substitution ciphers obscure the plain text. Encrypting plaintext and decrypting ciphertext is therefore calculated by

$$C_i \equiv m_i + k_i (\text{mod } 26), \quad m_i \equiv C_i - k_i (\text{mod } 26)$$

respectively, where subscript i denotes the $i-$th letter of the message m, key k, and cipher text C.

In 1854 the Vigenère cipher was cracked by Charles Babbage, by looking for repeated strings of three letters, known as trigrams, in the cipher text [11]. The key is likely to be a divisor of the amount of letters between two trigrams. Suppose that we have have found the trigram VSH repeated three times in a string of ciphertext; there are 12 letters between the first and second instance of the trigram and 15 between the second and third. It is possible that the same three letters of plain text were encrypted using the same three letters of the key, for example "the" encrypted to "VHS". The key length would therefore be a common divisor of 12 and 15 which suggests the key is 3 letters long.

For some time the Enigma machine used by Germany in World War II was thought be uncrackable. The Enigma machine uses a number of permutations and substitutions to encrypt individual letters. A simpler version of the Enigma machine was used commercially prior to the war, and in 1932 Marian Rejewski applied group theory, and the theory of permutations to decrypt messages [12].

It is evident that the history of cryptography is full of broken cryptosytems that were once thought secure, but how secure are our current systems? In order to understand the benefits of

Table 1: Vigenère Cipher

Plain text	t	h	e	q	u	i	c	k	b	r	o	w	n	f	o	x
Key	y	o	r	k	y	o	r	k	y	o	r	k	y	o	r	k
Cipher text	R	V	V	A	S	W	T	U	Z	F	F	G	L	T	F	H

quantum key distribution we must review the methods in use today, where they are beginning to fail and how QKD can solve these problems. We need to consider what makes a good cryptographic system and whether there is such a thing as a perfect system.

2.1 Symmetric Key Cryptography

In symmetric key cryptography Alice and Bob share a secret key K which is used to encrypt and decrypt messages. Most algorithms in use today are publicly available and therefore it is important that the ciphertext reveals little information about both the plain text and key. If the key is known messages can be decrypted, or if plain text is leaked into the cipher text this could reveal information about the key. There are two main types of symmetric key cryptography: stream ciphers, which encrypt character by character and block ciphers, which act on blocks of plain text. A symmetric-key encryption scheme consists of an encryption map $E : K \times M \to C$ such that for each $k \in K$ the map $E_k : M \to C; m \to E(k, m)$ is invertible. Here the elements $m \in M$ are plain text messages. C is the set of ciphertext and $k \in K$ are the keys. E_k is the encryption function with respect to the key k. The inverse function $D_k = E_k^{-1}$ is called the decryption function. Rijndael [1] , proposed by Daemen and Rijmen, is an example of a block cipher. The current advanced encryption standard (AES) is the Rijndael algorithm implemented with certain block sizes. We will not discuss the details of Rijndael but merely note its significance as a widely used symmetric-key encryption scheme. It is widely accepted that AES is secure (with correct implementation and adequate key sizes) and there have only been small incremental advances in methods to weaken the system. However, as history has taught us this does not necessarily mean Rijndael is secure and it is important to note that Rijndael has not been *proven* secure.

Vernam's One Time Pad

Vernam's one time pad is an example of a stream cipher. It is a *provably secure* cryptographic system unlike Rijndael [1]. It was C.E Shannon's contributions to information theory that were vital in proving the security of the one time pad. Shannon proved this by quantifiyng the amount of information an eavesdropper could gain from the ciphertext regarding the plain text[13]. The one time pad is said to be perfectly secret, if a random key is used no information regarding the plain text or key is leaked into the cipher text. Vernam's one time pad encrypts a binary string by bitwise XORing the string with a key. Consider a string $M := m_1m_2m_3...m_n$ where $m_i \in \{0, 1\}$ and a key stream $K := k_1k_2k_3...k_n$. Symmetric encryption and decryption are given by

$E(K, M) = K \oplus M = C$ and $D(K, C) = K \oplus C = M$, which is demonstrated below.

$$M \oplus K = 1010110010101101 \oplus 0101011010100100 = 1111101000001001 = C$$
$$C \oplus K = 1111101000001001 \oplus 0101011010100100 = 1010110010101101 = M$$

We can clearly see that the same key is used to encrypt and decrypt the message. The one time pad is significant because it is provably secure. One time pads are used to secure highly critical information however, there are two main problems with using them. First, generating a truly random string is difficult. This has been achieved in the past by observing radioactive decay. The second problem is an important one: distributing the key.

2.2 Asymmetric Key Cryptography

The cryptographic systems we have reviewed so far have the common problem of securely distributing keys. If Alice uses a shift cipher Bob needs to know how many places the letters are shifted to decrypt the key. If Alice uses the Vigènere cipher, AES or a one time pad Alice and Bob need to agree on an encryption key or exchange a codebook before transmitting the message. Key distribution can be achieved by exchanging the keys in person, through a secure delivery process, or by exchanging plain text over a public channel. All of these methods are either insecure or impractical.

Asymmetric key cryptography, often called public key cryptography provides a solution to the key exchange problem. In public key cryptography there are two keys: a public key used for encrypting messages and a private key that is kept secret and used for decryption. This is in contrast to symmetric key cryptography where the same key is used for encryption and decryption. Suppose that Alice has padlock and key and wants to send a secret message to Bob. Alice sends a box with the padlock to Bob. Bob puts his message in the box, fastens the padlock and sends it back to Alice. Alice has the key and therefore can unlock the box and read the message. This is analogous to the idea behind public key cryptography. More specifically, Alice has a *key pair* (k_p, k_s) where k_p is Alice's public key and k_s is Alice's private key. Given that k_p is shared publicly Bob can encrypt a message m yielding $C = E_{k_p}(k_p, m)$ where E is the encryption function. Alice decrypts the message using her secret key k_s and the decryption function D_{k_s} such that $m = D_{k_s}(k_s, C)$. To ensure this is considered secure enough for use, we require that computing m or the secret key k_s from $C = E_{k_p}(k_p, m)$ is a near impossible task.

2.2.1 RSA

RSA is an asymmetric key cryptosystem used to establish secure communication over public channels without having to physically exchange a key. RSA relies on the computationally hard problem of factorizing large prime numbers. We will briefly review the mathematics underpinning RSA.

Theorem 2.1 (Fermat's Little Theorem)**.** *Let p be a prime number and $a \in \mathbb{Z}$ be co-prime to p, then*

$$a^{p-1} \equiv 1 (\text{mod } p).$$

Table 2: RSA Summary

	Secret	Public
Alice	$p, q, d, \phi(n)$	n, e
Bob		C
Alice & Bob	M	

Proof. Consider the sequence of numbers $a, 2a, .., a(p-1)$. Then

$$a(2a)...(p-1)a \equiv 1(2)...(p-1)(\text{mod } p)$$

which simplifies to $a^{p-1}(p-1)! \equiv (p-1)!(\text{mod } p)$. By the cancellation law we have $a^{p-1} \equiv 1(\text{mod } p)$ as required. □

Definition 2.1 (Euler's Totient Function)**.** *Let $n \in \mathbb{Z}^+$. Then there are $\phi(n) \in \mathbb{Z}^+$ integers co-prime and not greater than n, such that*

$$\phi(n) = n\Pi_{p|n}(1 - \frac{1}{p}),$$

where p are prime divisors of n.

Theorem 2.2 (Euler's Theorem)**.** *If a and n are co-prime integers then*

$$a^{\phi(n)} \equiv 1(\text{mod } n).$$

Proof. Let A be the set of positive integers co-prime and not greater than n, i.e

$$A = \{c \leq n : (n, c) = 1\},$$

and $|A| = \phi(n)$. Let $a \in \mathbb{Z}$ and $c_i \in A$ for $i = 1, 2, ...\phi(n)$. Then

$$a^{\phi(n)}c_1c_2...c_{\phi(n)} \equiv ac_1ac_2...ac_{\phi(n)}(\text{mod } n),$$

and thus by the cancellation law $a^{\phi(n)} = 1(\text{mod } n)$. □

The algorithm to implement RSA goes as follows. Alice chooses two large distinct primes p, q and calculates $n = pq$. Alice publicly shares n and keeps p, q secret. Alice then publishes an encryption exponent e such that $(\phi(n), e) = 1$, and calculates the decryption exponent d such that $ed \equiv 1(\text{mod } \phi(n))$ and keeps d secret. Bob encrypts a message M by using Alice's public encryption exponent e and calculating $C \equiv M^e(\text{mod } n)$. Bob sends the encrypted message C to Alice. Alice then decrypts C by calculating $C^d \equiv M^{ed} \equiv M^{1+c\phi(n)} \equiv M(M^c)^{\psi(n)} \equiv M(\text{mod } n)$.

2.2.2 The Integer Factoring Problem

The security of RSA is reliant on the difficult task of decomposing large numbers into their prime factors. There have been a number of advancements in this task over the past few decades, some of which have been motivated by RSA factoring challenges. The challenge goes as follows: given the integer n there exists two prime numbers p and q where $n = pq$, what are p and q?

This problem can be solved inefficiently by brute force methods, or instead clever techniques can be applied. The most effective method in use today is the number field sieve, which saw a significant reduction in the time taken to factor large primes compared to the quadratic sieve. For further information regarding these methods read [1]. In 1994 Peter Shor proposed a quantum algorithm [3] for implementation on a quantum computer, that can be used to decompose an integer into its prime factors. Shor's algorithm can factor large primes in polynomial time which is a significant improvement on the generalised number field sieve that works in sub-exponential time. This significant reduction in computational time means that even larger prime numbers would need to be used in RSA to ensure secure encryption, or perhaps we should consider different methods of encryption altogether. Finding large prime numbers for use in cryptography is a problem in itself and there are organisations and challenges that are devoted to finding such numbers.

2.3 What next for key distribution?

Distributing keys securely is a common problem in cryptography and it underpins the majority of encrypted communication over public channels such as the internet. Looking to the future, we need alternate methods of distributing keys to establish secure lines of communications. We have seen throughout history that our cryptographic methods have been broken and a fully functional quantum computer capable of implementing Shor's algorithm at scale will further exacerbate this issue. Fortunately, quantum phenomena offers us a solution. Quantum key distribution can be used to implement unconditionally secure protocols such as the perfectly secure one time pad. There are a number of reasons why QKD, and in particular entanglement-based QKD, are preferable to classical methods. We will discuss advantages such as the No-Cloning Theorem, the ability to generate truly random keys and the ability to detect eavesdropping in the following sections.

3 Mathematical Preliminaries

3.1 Hilbert Spaces, Operators and Tensor Products

The mathematical foundation of quantum mechanics relies on Hilbert spaces and operators. A Hilbert space is a vector space over $\mathbb{C}$ with an inner product and the property of completeness with respect to the norm metric; for our purposes we will only be concerned with finite dimensional Hilbert spaces. It can be shown that all finite dimensional complex inner product spaces are complete, we will briefly review some important mathematical background regarding Hilbert spaces; a more complete description can be found in [14].

Definition 3.1 (Inner Product)**.** *Let $\mathcal{H}$ be a complex vector space with a complex-valued function $\langle\cdot|\cdot\rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$, then $\langle\cdot|\cdot\rangle$ is an inner product if:*

$$\langle\varphi|c\psi + \phi\rangle = c\,\langle\varphi|\psi\rangle + \langle\varphi|\phi\rangle\,, \tag{1}$$

$$\langle\varphi|\phi\rangle^* = \langle\phi|\varphi\rangle\,, \tag{2}$$

$$\langle\psi|\psi\rangle \geq 0, \quad \langle\psi|\psi\rangle = 0 \iff \psi = 0, \tag{3}$$

for all $\psi, \phi, \varphi \in \mathcal{H}$ and $c \in \mathbb{C}$.

A complex vector space with an inner product is called an inner product space. Inner products are useful for giving vector spaces a convenient geometrical interpretation providing notions such as the length of a vector and the angle between two vectors. Two vectors ψ and ϕ if $\langle\phi|\psi\rangle = 0$ then $|\psi\rangle$ and $|\phi\rangle$ are orthogonal.

Definition 3.2. *Let $\mathcal{H}$ be a complex vector space. Then a function $||\cdot|| : \mathcal{H} \to \mathbb{R}$ is a norm if:*

$$||\varphi|| \geq 0,\ ||\varphi|| = 0 \iff \varphi = 0,$$

$$||c\varphi|| = |c|||\varphi||,$$

$$||\varphi + \psi|| \leq ||\varphi|| + ||\psi||,$$

for all $\varphi, \psi \in \mathcal{H}$ and $c \in \mathbb{C}$.

In quantum mechanics the natural norm to use is $|\,\langle\psi|\psi\rangle\,| = \sqrt{\langle\psi|\psi\rangle}$. It is fairly common for single bars to be used instead of double bars as in definition 3.2. It can be shown that every inner product space is a normed space with the norm $||\psi|| = \sqrt{\langle\psi|\psi\rangle}$. Norms are useful because they give vector spaces a notion of distance. The distance between two vectors ψ, ϕ can be defined as $d(\psi, \phi) = ||\psi - \phi||$, the function d is known as a metric and therefore the norm is said to induce a metric on $\mathcal{H}$.

Definition 3.3. *A* metric space *is a set $X \neq \emptyset$ with a metric $d : X \times X \mapsto \mathbb{R}$ often denoted as the ordered pair (X, d) with the following properties $\forall x, y, z \in X$:*

$$d(x, y) \geq 0,\ d(x, y) = 0 \iff x = y$$

$$d(x, y) = d(y, x)$$

$$d(x, y) \leq d(x, z) + d(z, y)$$

The metric d is referred to as the distance function in some literature, if we think of d as measuring some form of distance we can make sense of these properties intuitively. The first property states that distances are non-negative, and additionally that the distance $d(x, x) = 0$ i.e measuring the distance from a point to itself is 0. The second states that the distance between two points is the same regardless of which way you measure it. Lastly, the third point is the triangle inequality, the distance is always shorter measuring directly from x to y as opposed to measuring via some third point denoted z [16].

Definition 3.4. *A sequence $\{\varphi_j\}$ is a Cauchy sequence if, for every $\epsilon > 0$, there exists a positive integer N_ϵ such that $d(\varphi_j, \varphi_k) < \epsilon$ whenever $j, k > N_\epsilon$.*

Definition 3.5. *A metric space (X, d) is called complete if every Cauchy sequence in X converges to an element of X.*

An example of an incomplete metric space is the space $(0, 1)$ with the metric $d(x, y) = |x - y|$. Consider the Cauchy sequence $x_n = \frac{1}{x_n}$, then $\lim_{n\to\infty} x_n \to 0 \notin X$. Then, by definition $((0, 1), d)$ is not a complete metric space.

Definition 3.6 (Hilbert Space)**.** *A complete inner product space is called a Hilbert space.*

As stated earlier, we are only concerned with finite dimensional Hilbert space. The space $\mathbb{C}^n$ with an inner product has the property of completeness with respect to the norm metric and is therefore a Hilbert space. Without dwelling on the details this can be seen by proving that $\mathbb{R}$ is a complete metric space and showing that $\mathbb{R}^2$ is isomorphic to $\mathbb{C}$. This can be generalised to $\mathbb{C}$. In this report we will refer to linear operators as operators and will identify operators as matrices in the set $M_{d\times d}(\mathbb{C})$.

Definition 3.7 (Unitary Operator)**.** *An operator U is unitary if it satisfies $UU^\dagger = U^\dagger U = \mathbb{I}$.*

Unitary operators are useful because they allow us to change between bases. That is for two orthonormal bases $\{|i\rangle\}$ and $\{|\hat{i}\rangle\}$ there exists a unitary operator U such that $U|i\rangle = |\hat{i}\rangle$.

Definition 3.8. *An operator A is said to be positive if $\langle\psi| A |\psi\rangle \geq 0$, $\forall |\psi\rangle \in \mathcal{H}$.*

An equivalent characterisation is that all of the eigenvalues $\{\lambda_i\}$ of A are such that $\lambda_i \geq 0$.

Definition 3.9 (Tensor Product)**.** *Let V and W be vector spaces with dimensions n and m (we will assume that V and W are Hilbert spaces). The tensor product, $V \otimes W$, is a vector space with dimensions nm.*

The elements of $V \otimes W$ are the linear combininations of $|v\rangle \otimes |w\rangle$ for $|v\rangle \in V$ and $|w\rangle \in W$. The tensor product is a way of putting vector spaces together to form larger vector spaces and is the standard tool for understanding composite quantum systems. We do not need to concern ourselves with the intracacies of the tensor product, for our purposes we will use an explicit form for vectors and matrices which is shown in the examples below. A more detailed description can be found in [15].

Example 3.1. *Consider two state vectors, $|v\rangle \in \mathcal{H}_1$ and $|w\rangle \in \mathcal{H}_2$ where the dimensions of $\mathcal{H}_1$ and $\mathcal{H}_2$ are n and m respectively. We can write the two states explicitly as*

$$|v\rangle = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \ |w\rangle = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix},$$

then the tensor product $|v\rangle \otimes |w\rangle$ is

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \otimes |w\rangle = \begin{pmatrix} v_1 |w\rangle \\ v_2 |w\rangle \\ \vdots \\ v_n |w\rangle \end{pmatrix} = \begin{pmatrix} v_1 w_1 \\ v_1 w_2 \\ \vdots \\ v_1 w_m \\ v_2 w_1 \\ v_2 w_2 \\ \vdots \\ v_n w_m \end{pmatrix} \in \mathbb{C}^{nm}.$$

Let $|v_1\rangle, |v_2\rangle \in V$, $|w_1\rangle, |w_2\rangle \in W$, and $z \in \mathbb{C}$. By definition $V \otimes W$ has the following properties:

$$\begin{aligned} z(|v\rangle \otimes |w\rangle) &= (z\,|v\rangle) \otimes |w\rangle = |v\rangle \otimes (z\,|w\rangle), \\ (|v_1\rangle + |v_2\rangle) \otimes |w\rangle &= |v_1\rangle \otimes |w\rangle + |v_2\rangle \otimes |w\rangle, \\ |v\rangle \otimes (|w_1\rangle + |w_2\rangle) &= |v\rangle \otimes |w_1\rangle + |v\rangle \otimes |w_2\rangle \end{aligned}$$

These basic properties will be used throughout this report. An inner product can naturally be defined on $V \otimes W$ such that $\langle v_1 \otimes w_1 | v_2 \otimes w_2 \rangle = \langle v_1 | v_2 \rangle \langle w_1 | w_2 \rangle$. Next we will show the explicit form for the tensor product between two matrices. It is convenient to use the matrix representation of an operator, the operation $\otimes$ between two matrices is known as the *Kronecker product*.

Example 3.2. *Let $A \in M_n(\mathbb{C})$ be an $n \times n$ matrix and $B \in M_m(\mathbb{C})$ be an $m \times m$ matrix. The matrix representation of $A \otimes B$ is an $nm \times nm$ matrix and $A \otimes B \in M_{nm}(\mathbb{C})$.*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mm} \end{pmatrix}.$$

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \dots & a_{nn}B \end{pmatrix}.$$

It can be shown that operators act on kets such that $(A \otimes B)\,|v\rangle \otimes |w\rangle = A\,|v\rangle \otimes B\,|w\rangle$. Calculating this expicitly for finite square matrices is laborious, instead we will verify $(A \otimes \mathbb{I}_3)\,|v\rangle \otimes |w\rangle = A\,|v\rangle \otimes |w\rangle$ with $A \in M_2(\mathbb{C})$ and $\mathbb{I} \in M_3(\mathbb{C})$.

Example 3.3.

$$(A \otimes \mathbb{I})\,|v\rangle \otimes |w\rangle = \begin{pmatrix} a_{11} & 0 & 0 & a_{12} & 0 & 0 \\ 0 & a_{11} & 0 & 0 & a_{12} & 0 \\ 0 & 0 & a_{11} & 0 & 0 & a_{12} \\ a_{21} & 0 & 0 & a_{22} & 0 & 0 \\ 0 & a_{21} & 0 & 0 & a_{22} & 0 \\ 0 & 0 & a_{21} & 0 & 0 & a_{22} \end{pmatrix} \begin{pmatrix} v_1w_1 \\ v_1w_2 \\ v_1w_3 \\ v_2w_1 \\ v_2w_2 \\ v_2w_3 \end{pmatrix} = \begin{pmatrix} (a_{11}v_1 + a_{12}v_2)w_1 \\ (a_{11}v_1 + a_{12}v_2)w_2 \\ (a_{11}v_1 + a_{12}v_2)w_3 \\ (a_{21}v_1 + a_{22}v_2)w_1 \\ (a_{21}v_1 + a_{22}v_2)w_2 \\ (a_{21}v_1 + a_{22}v_2)w_3 \end{pmatrix}$$

which equals $A\,|v\rangle \otimes |w\rangle$ as required.

Similarly one can calculate $(\mathbb{I} \otimes B)\,|v\rangle \otimes |w\rangle = |v\rangle \otimes B\,|w\rangle$ and use that $(A \otimes \mathbb{I})(\mathbb{I} \otimes B) = A \otimes B$ to see that $(A \otimes B)\,|v\rangle \otimes |w\rangle = A\,|v\rangle \otimes B\,|w\rangle$. Lastly, note that we can naturally write

$$\begin{aligned}(A_1\,|v_1\rangle \otimes B_1\,|w_1\rangle)^\dagger)(A_2\,|v_2\rangle \otimes B_2\,|w_2\rangle) &= (\langle v_1|\,A_1^\dagger \otimes \langle w_1|\,B_1^\dagger)(A_2\,|v_2\rangle \otimes B_2\,|w_2\rangle) \\ &= \langle v_1|\,A_1^\dagger A_2\,|v_2\rangle\,\langle w_1|\,B_1^\dagger B_2\,|w_2\rangle\,.\end{aligned}$$

3.2 Quantum Mechanics

In quantum mechanics, a state refers to the state of a quantum system. Using Dirac's Bra-Ket notation, states are denoted by *kets* $|\psi\rangle \in \mathcal{H}$ which are rays in Hilbert space and can be represented by vectors. A *bra* $\langle\psi| \in \mathcal{H}$ is the adjoint of the ket, that is $|\psi\rangle^\dagger = \langle\psi|$.

Definition 3.10. *A pure state $|\psi\rangle \in \mathcal{H}$ is a state that can be written as a unit vector, that is $\langle\psi|\psi\rangle = 1$. If a state is not pure then we say it is a mixed state.*

Postulates are the set of rules quantum mechanics is built upon. The postulates for pure states are as follows.

Postulates for Pure States

1. For each physical system there is an associated Hilbert space denoted $\mathcal{H}$. Any pure state of a system can be represented as an element $|\psi\rangle \in \mathcal{H}$, such that $\langle\psi|\psi\rangle = 1$

2. The evolution of a closed quantum system is described by a unitary transformation $U : \mathcal{H} \to \mathcal{H}$. That is, the state of a system $|\psi\rangle$ at time t_1 is related to the state $|\psi'\rangle$ at time t_2 by a unitary operator U which depends only on the times t_1 and t_2, such that $|\psi'\rangle = U\,|\psi\rangle$.

3. A measurement on a physical system with Hilbert space $\mathcal{H}$ is represented by a collection of measurement operators, $\{M_x\}$. These are linear operators, $M_x : \mathcal{H} \to \mathcal{H}$ that satisfy $\sum_x M_x^\dagger M_x = \mathbb{I}$.

4. If a system with state $|\psi\rangle \in \mathcal{H}$ is measured with measurement $\{M_x\}$, then the probability of outcome $X = x$, such that X is a random variable is $P_X(x) = \langle\psi| M_x^\dagger M_x |\psi\rangle$, and the post-measurement state on obtaining $X = x$ is

$$|\psi_x\rangle = \frac{M_x |\psi\rangle}{\langle\psi| (M_x M_x^\dagger) |\psi\rangle}.$$

5. If we have a system with state space $\mathcal{H}_1$ and a system with state space $\mathcal{H}_2$, then the state space of the joint system is the tensor product of these Hilbert spaces, $\mathcal{H}_1 \otimes \mathcal{H}_2$. If the first system is in the state $|\psi\rangle_1 \in \mathcal{H}_1$ and the second is in the state $|\psi\rangle_2 \in \mathcal{H}_2$, then the joint system is in the *product state* $|\psi\rangle_1 \otimes |\psi\rangle_2 \in \mathcal{H}_1 \otimes \mathcal{H}_2$.

Definition 3.11 (POVM)**.** *A positive operator valued measure (POVM) is a collection of positive operators $\{E_x\}$ such that $\sum_x E_x = \mathbb{I}$.*

Note that POVMs are related to measurement operators in that $E_x = M_x^\dagger M_x$. Sometimes the measurement operator associated with E_x is simply denoted $\sqrt{E_x}$, note that this is not a unique choice. Recall that measurements can also be made in terms of observables, where observables refer to Hermitian operators.

Definition 3.12. *The Pauli matrices are a Hermitian, unitary, set of three 2×2 matrices which are denoted by $\sigma_x, \sigma_y, \sigma_z$ and defined below:*

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

These three matrices, along with the identity matrix denoted $\mathbb{I}$ (often denoted σ_0 in this context) form an orthonormal basis and a 4-dimensional complex vector space; each matrix corresponds to an angular momentum observable. Importantly, the Pauli spin operators are used to describe spin measurements of particles which will be explained in section 5.1. In the following definition $\mathcal{L}(\mathcal{H})$ denotes the set of linear operators on $\mathcal{H}$.

Definition 3.13. *The trace of $A \in \mathcal{L}(\mathcal{H})$ denoted $tr(A)$ is defined by*

$$tr : \mathcal{L}(\mathcal{H}) \to \mathbb{C}; A \mapsto \sum_i \langle i| A |i\rangle,$$

where $\{|i\rangle\}$ form an orthonormal basis of $\mathcal{H}$.

A problem that we will encounter in section 5 is distinguishing between two non-orthogonal states. If we have two non-orthogonal states $|\psi_1\rangle$ and and $|\psi_2\rangle$ it is not possible to perform a measurement to distinguish between the two states with certainty. We now present this problem, as may be seen in [17, 15].

Lemma 3.1. *Let $|\psi_1\rangle \in \mathcal{H}$ and $|\psi_2\rangle \in \mathcal{H}$ be two non-orthogonal states, then there is no measurement that can distinguish with complete certainty between the two states.*

Proof. Suppose that we can take measurements such that we can distinguish between the two states $|\psi_1\rangle$ and $|\psi_2\rangle$ with complete certainty, then we require that $\langle\psi_1| E_1 |\psi_1\rangle = 1$ and $\langle\psi_2| E_2 |\psi_2\rangle = 1$. Since $\sum_i E_i = \mathbb{I}$ we have by definition $\sum_i \langle\psi_1| E_i |\psi_1\rangle = 1$ which implies that $\langle\psi_1| E_2 |\psi_1\rangle = 0$ and thus $\sqrt{E_2}|\psi_1\rangle = 0$. Since $|\psi_1\rangle$ and $|\psi_2\rangle$ are not orthogonal we can decompose $|\psi_2\rangle$ such that $|\psi_2\rangle = \alpha|\psi_1\rangle + \beta|\varphi\rangle$, where $|\varphi\rangle$ is orthonormal to $|\psi_1\rangle$ and $|\alpha|^2 + |\beta|^2 = 1$ and $|\beta| < 1$. Then $\sqrt{E_2}|\psi_2\rangle = \beta|\varphi\rangle$, which implies that $\langle\psi_2| E_2 |\psi_2\rangle = |\beta|^2 \langle\varphi| E_2 |\varphi\rangle \leq |\beta|^2 < 1$ which is a contradiction. □

Example 3.4 (Distinguishing between non-orthogonal states with POVMs)**.** *Consider two states $|\psi_1\rangle = |0\rangle$ and $|\psi_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Note that these states are not orthogonal as $\langle\psi_1|\psi_2\rangle \neq 0$ and consider the POVM*

$$E_1 = \frac{\sqrt{2}}{1+\sqrt{2}}|1\rangle\langle 1|, \quad E_2 = \frac{\sqrt{2}}{2(1+\sqrt{2})}(|0\rangle - |1\rangle)(\langle 0| - \langle 1|), \quad E_3 = \mathbb{I} - E_1 - E_2.$$

The the expected values are

$$\langle\psi_1| E_1 |\psi_1\rangle = 0, \quad \langle\psi_1| E_2 |\psi_1\rangle = \frac{\sqrt{2}}{2(1+\sqrt{2})}, \quad \langle\psi_1| E_3 |\psi_1\rangle = \frac{1}{\sqrt{2}},$$

$$\langle\psi_2| E_1 |\psi_2\rangle = \frac{\sqrt{2}}{2(1+\sqrt{2})}, \quad \langle\psi_2| E_2 |\psi_2\rangle = 0, \quad \langle\psi_2| E_3 |\psi_2\rangle = \frac{1}{\sqrt{2}}.$$

From these measurements we can see that if the result is E_1 the state cannot be $|\psi_1\rangle$ and similarly if the result is E_2 then the state cannot be $|\psi_2\rangle$. We can therefore distinguish between the two non-orthogonal states but the probability of an inconclusive result for either state is $\frac{1}{\sqrt{2}}$ and therefore we do not have complete certainty.

Quantum mechanics is often described using state vectors. This is the natural language to describe pure states, however, there is an equivalent formulation using the density operator. Density operators can be used to describe both mixed and pure states.

Definition 3.14. *A mixed state is a statistical ensemble of states with weighted probabilities and cannot be written as a single ket or pure state.*

It follows that density operators are a generalisation of the bra-ket formulation of quantum mechanics stated earlier. Density operators are more convenient for describing complex quantum systems.

Definition 3.15 (Density Operator)**.** *Given a finite statistical ensemble $\{|\psi_i\rangle, p_i\}$ of n pure states $|\psi_i\rangle$ with associated probabilities $p_i \geq 0$ the* density operator *ρ is defined*

$$\rho = \sum_i^n p_i |\psi_i\rangle\langle\psi_i|, \quad \sum_i^n p_i = 1.$$

The density operator can also be characterised as a positive operator with $\mathrm{tr}\rho = 1$. Consider the state $|\Phi\rangle \in \mathcal{H}$ then we can write $\langle\Phi|\,\rho\,|\Phi\rangle = \langle\Phi|\,(\sum_i p_i\,|\psi_i\rangle\,\langle\psi_i|)\,|\Phi\rangle$ such that $\langle\Phi|\,\rho\,|\Phi\rangle = \sum_i p_i\,\langle\Phi|\psi_i\rangle\,\langle\psi_i|\Phi\rangle = \sum_i p_i|\,\langle\psi|\Phi_i\rangle\,|^2 \geq 0$ since $p_i \geq 0$, and thus ρ is positive. Now consider an orthonormal basis $\{|j\rangle\}$ for $\mathcal{H}$, then $tr\rho = \sum_{i,j}^n p_i\,\langle j|\psi_i\rangle\,\langle\psi_i|j\rangle = \sum_{i,j}^n p_i\,\langle\psi_i|\,(|j\rangle\,\langle j|)\,|\psi_i\rangle = \sum_i^n p_i\,\langle\psi_i|\psi_i\rangle = 1$ as required. Note that a pure state $\rho = |\psi\rangle\,\langle\psi|$ has the property that $\mathrm{tr}\rho^2 = \langle\psi|\psi\rangle = 1$. Density operators may also be represented as a mixture of other density operators with weighted probabilities p_i, that is $\rho = \sum_i p_i\rho_i$.

Postulates for Density Operators

1 For each physical system there is an associated Hilbert space, $\mathcal{H}$. Any general, mixed or pure state of a system can be represented by a density operator ρ, defined as a positive operator with trace 1. The set of density operators on a Hilbert space $\mathcal{H}$ is denoted by $S(\mathcal{H})$.

2 The evolution of a closed quantum system is described by a unitary transformation $U : \mathcal{H} \to \mathcal{H}$. That is, the state of a system ρ at time t_1 is related to the state ρ' at time t_2 by a unitary operator U which depends only on the times t_1 and t_2, such that $\rho' = U\rho U^\dagger$.

2 A measurement on a physical system is represented by a collection of measurement operators, $\{M_x\}$. These are linear operators, $M_x : \mathcal{H}_1 \to \mathcal{H}_2$ that satisfy $\sum_x M_x^\dagger M_x = \mathbb{I}$.

3 If a system with density operator $\rho \in S(\mathcal{H}_1)$ is measured with measurement $\{M_x\}$, then the probability of outcome x is $P_X(x) = tr(M_x^\dagger M_x\rho)$, and the post-measurement state on obtaining $X = x$ is

$$\rho_x = \frac{M_x\rho M_x^\dagger}{\mathrm{tr}(M_x\rho M_x^\dagger)}.$$

4 If we have a system with state space $S(\mathcal{H}_1)$ and a second system with state space $S(\mathcal{H}_2)$, then the state space of the joint system is the tensor product of these Hilbert spaces, $S(\mathcal{H}_1 \otimes \mathcal{H}_2)$. If the first system is in the state $\rho \in S(\mathcal{H}_1)$ and the second is in the state $\sigma \in S(\mathcal{H}_2)$, then the joint system is in the *product state* $\rho \otimes \sigma \in S(\mathcal{H}_1 \otimes \mathcal{H}_2)$. This naturally generalises to more spaces.

Definition 3.16. *Let A be a linear operator such that $A : \mathcal{H}_1 \otimes \mathcal{H}_2 \to \mathcal{H}_1 \otimes \mathcal{H}_2$. Then the partial trace of A over $\mathcal{H}_1$ is defined by the mapping $tr_1 : \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2) \to \mathcal{L}(\mathcal{H}_2)$ such that*

$$tr_1 A = \sum_i (\langle i|_1 \otimes \mathbb{I}_2)A(|i\rangle_1 \otimes \mathbb{I}_2),$$

where $\{|i\rangle\}$ and $\{|\hat{i}\rangle\}$ are orthonormal bases for $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively.

The partial trace is a convenient tool for dealing with subsystems of an overall system, this is applied in section 5.

3.3 Quantum Entanglement

We will now mathematically define quantum entanglement. Information regarding the history, implications and applications of entanglement can be found in section 4 and 5. So far we have touched upon pure states, mixed states and product states, two other types of states that are integral to the theory of entanglement are *separable states* and *entangled states*.

Definition 3.17. *Let A and B be two quantum systems with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively, and let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. A* pure state *$|\psi\rangle \in \mathcal{H}$ is called* separable *if and only if $|\psi\rangle = |\psi\rangle_A \otimes |\psi\rangle_B$, for some $|\psi\rangle_A \in \mathcal{H}_A$ and $|\psi\rangle_B \in \mathcal{H}_B$, otherwise $|\psi\rangle$ is called* entangled.

Four important maximally entangled states are the *Bell states*, also known as *EPR pairs*. The Bell states are $|\phi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$ and $|\psi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$. In the example below we will show that $|\phi^+\rangle$ is an entangled state.

Example 3.5. *Suppose that the $|\phi^+\rangle$ is a separable state, that is we can write $|\phi^+\rangle$ such that $|\phi\rangle^+ = |\psi\rangle_A \otimes |\psi\rangle_B$ for some $|\psi\rangle_A \in \mathcal{H}_A$ and $|\psi\rangle_B \in \mathcal{H}_B$. Consider the normalised states $|\psi\rangle_A = \frac{1}{\sqrt{2}}(a|0\rangle + b|1\rangle)$ and $|\psi\rangle_B = \frac{1}{\sqrt{2}}(\alpha|0\rangle + \beta|1\rangle)$. Then*

$$|\psi\rangle_A \otimes |\psi\rangle_B = a\alpha|0\rangle \otimes |0\rangle + b\beta|1\rangle \otimes |1\rangle + a\beta|0\rangle \otimes |1\rangle + \alpha b|1\rangle \otimes |0\rangle .$$

We require that the coefficients of the third and fourth term are zero which leads to a contradiction. If $a\beta = 0$ this implies that either $a = 0$ or $\beta = 0$. If we choose $a = 0$ then the first term vanishes and if we choose $\beta = 0$ the second term vanishes. A similar argument can be made for the condition that $\alpha b = 0$. Thus we conclude that $|\phi^+\rangle$ is an entangled state.

Definition 3.18. *Let A and B be two quantum systems with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively, and let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ for some density operator $\rho \in \mathcal{S}(\mathcal{H})$. Then ρ is called* separable *if and only if $\rho = \sum_i p_i \rho_i^A \otimes \rho_i^B$, for $\rho_i^A \in \mathcal{S}(\mathcal{H}_A)$ and $\rho_i^B \in \mathcal{S}(\mathcal{H}_B)$, otherwise ρ is called* entangled.

3.3.1 Positive Partial Transpose

Checking whether a mixed state is separable or entangled is not as simple as checking whether a pure state is, there are a number of clever ways of testing for entanglement and this is an active area of research. Whilst there are a number of techniques to test for entanglement we will use the positive partial transpose criterion also known as the Peres-Horodecki criterion [18]. The test is easy to perform and can clearly be applied to both pure states and mixed states, we will use this result in 5.

Definition 3.19 (Partial Transpose)**.** *Let $\{|i\rangle\}$ and $\{|\hat{i}\rangle\}$ be orthonormal bases for $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. Then $\rho = \sum_{i,j,k,l} \rho_{i,j,k,l} |i\rangle\langle j| \otimes |k\rangle\langle l|$, with elements $\rho_{i,j,k,l} = (\langle i| \otimes \langle k|)\rho(|j\rangle \otimes |l\rangle)$. The* partial transpose *of ρ with respect to the A subsystem is*

$$\rho^{T_A} = \sum_{i,j,k,l} \rho_{i,j,k,l} (|i\rangle\langle j|)^T \otimes |k\rangle|l\rangle = \sum_{i,j,k,l} \rho_{j,i,k,l} |j\rangle\langle i| \otimes |k\rangle\langle l| .$$

The following lemma, known as either the Peres-Horodecki criterion or Positive Partial Transpose lemma, can be used to establish whether a density operator is an entangled state [18]. Recall that any pure state can be written as a density operator such that $\rho = |\psi\rangle\langle\psi|$.

Lemma 3.2 (Positive Partial Transpose)**.** *If $\rho_{AB} \in S(\mathcal{H}_A \otimes \mathcal{H}_B)$ is separable then $\rho_{AB}^{T_B}$ is a positive operator.*

Proof. Let ρ_{AB} be a density operator of the form $\rho_{AB} = \sum_{i,j,k,l} \rho_{i,j,l,k} |i\rangle\langle j| \otimes |k\rangle\langle l|$. Then taking the partial transpose over the B subsystem yields

$$\rho^{T_B} = \sum_{i,j,k,l} \rho_{i,j,k,l} |i\rangle\langle j| \otimes (|k\rangle\langle l|)^T = \sum_{i,j,k,l} \rho_{i,j,l,k} |i\rangle\langle j| \otimes (|l\rangle\langle k|).$$

By the assumption that ρ is separable and by definition 3.18, ρ can be decomposed into density operators of the two subsystems such that $\rho_{\text{sep}} = \sum_i p_i(\rho_i^A) \otimes (\rho_i^B)^T$. Now, suppressing the subscript, clearly $\text{tr}\rho^B = \text{tr}(\rho^B)^T = 1$. Since ρ^B is a valid density operator, it is a positive operator. Thus $\langle\psi|\rho^B|\psi\rangle = \sum_{k,l} \rho_{k,l}^B |k\rangle\langle l| \geq 0$ and $\langle\psi|(\rho^B)^T|\psi\rangle = \sum_{k,l} \rho_{k,l}^B |l\rangle\langle k| = \sum_{k,l} \rho_{l,k}^B |k\rangle\langle l| \geq 0$. It follows that $(\rho^B)^T$ is a positive operator and therefore, the partial transpose of ρ_{AB} is positive and has trace 1. Thus ρ_{AB} is a separable state. □

Note that $\rho^{T_A} = (\rho^{T_B})^T$, that is $(\rho^{T_B})^T = \sum_{i,j,k,l} \rho_{i,j,l,k} |j\rangle\langle i| \otimes |k\rangle\langle l| = \rho^{T_A}$, thus the test does not depend on which subsystem is transposed.

Example 3.6. *Consider the Bell state $|\phi^+\rangle$, which we can write as a density operator*

$$\rho = |\phi^+\rangle\langle\phi^+| = \frac{1}{2}(|00\rangle + |11\rangle)(\langle 00| + \langle 11|).$$

Expanding the brackets yields terms of the form $|ii\rangle\langle jj|$ which can be written more explicitly in the form used in 3.2 as $(|i\rangle \otimes |i\rangle)(\langle j| \otimes \langle j|) = |i\rangle\langle j| \otimes |i\rangle\langle j|$ Taking the partial transpose over the B system yields

$$\rho^{T_B} = |0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1| + |0\rangle\langle 1| \otimes (|0\rangle\langle 1|)^T + |1\rangle\langle 0| \otimes (|1\rangle\langle 0|)^T$$

which is also written in matrix form below:

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \quad \rho^{T_A} = \rho^{T_B} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Calculating the characteristic polynomial we find that $\det(\rho^{T_B} - \lambda\mathbb{I}) = (2\lambda - 1)^3(2\lambda + 1) = 0$, with eigenvalues $\lambda = \frac{1}{2}$ and $\lambda = -\frac{1}{2}$. Since one eigenvalue is negative, by the positive partial transpose criterion ρ is entangled.

4 EPR, Bell's Theorem and Entanglement

In the famous Einstein-Podolsky-Rosen paper published in 1935, Einstein et al. [7] argued that entanglement phenomena made quantum theory incomplete. They thought that it should be possible to predict an outcome before a measurement is taken with complete certainty and that entanglement phenemona contradicted this idea [15]. In classical mechanics probabilistic models tend to be based on an assumption that a system has a definite, or predetermined value before we measure it. Probabilistic models can be used in cases where we lack complete knowledge of the system. This may be because it is either infeasible to model the entire system or there is some unknown or hidden variable not included in the model.

At the heart of the EPR paradox are the assumptions of reality and locality. The assumption of *reality* refers to the idea that physical properties have definite values which exist independent of observation. The assumption of *locality* refers to the idea that reality in one location is not affected by measurements performed simultaneously in a distant location [15]. Einstein et al. considered whether entanglement phenemona had a physical counterpart and could be explained by a local hidden variable theory [19].

4.1 Bohm's Simplification

Entanglement phenemona is best demonstrated by considering a simpler argument presented by David Bohm in 1951 [20]. Bohm considered a spin-zero particle that decays into two spin-$\frac{1}{2}$ particles. Consider the entangled state

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |1\rangle_B - |1\rangle_A \otimes |0\rangle_B),$$

here the subscripts label the particles with an associated Hilbert space. If we measure the spin of particle A and find it to be in the state $|0\rangle$ then the states of *both* particles in the system instantaneously collapse to $|0\rangle_A \otimes |1\rangle_B$. Similarly if we measure the state of particle A and find the state $|1\rangle_A$ the state of both particles instantaneously collapses to $|1\rangle_A \otimes |0\rangle_B$. From this argument it is clear that there is a correlation between the two particles.

4.2 Bell's Theorem and the CHSH Inequality

In 1964 John Stewart Bell published a paper which approached the issue of hidden variables [8]. Bell considered Bohm's version of the EPR experiment and showed that statistical predictions of any hidden variable theory which satisfies the principle of locality and reality must be bounded by what is now referred to as a *Bell inequality* [21]. Bell inequalities deduce upper and lower bounds on the statistical correlations of measurement outcomes in theories that assume locality and reality. In 1969 Clauser, Horne, Shimony and Holt proposed a generalised Bell inequality now known as the CHSH inequality [9]. The original Bell inequality is restricted to the case where measurement results are completely anti-correlated; the CHSH inequality is not restricted to this case, is easier to derive and can be used experimentally.

We will follow Bell's 1971 derivation [22] of the CHSH inequality. First, suppose that Alice and Bob are separated by a distance ct, where c is the speed of light and $t = 1$. A third party, Charlie, prepares two particles and sends one to Alice and one to Bob. Alice can measure one of two properties, P_A or $P_{A'}$ and similarly Bob can measure one of two properties P_B or $P_{B'}$. The measurements of each properties have values $A = \pm 1$, $A' = \pm 1$, $B = \pm 1$, $B' = \pm 1$ respectively. Alice and Bob measure the two properties simultaneously and randomly choose which property to measure. The simultaneous measurements mean that Alice's measurement cannot disturb Bob's and Bob's cannot disturb Alice's.

Under the assumption that a local hidden variable λ exists, suppose that the complete description of the initial state is dependent on a hidden variable with probability distribution $\rho(\lambda)$. Given measurement settings a and b, A is dependent on both a and λ and therefore we can write $\hat{A}(a, \lambda)$ and similarly $\hat{B}(b, \lambda)$. Under the assumption that locality holds we require that $\hat{A}$ does not depend on $\hat{B}$ or vice-versa. The expected value of the product $\hat{A}\hat{B}$ is then

$$E(a, b) = \int \hat{A}(a, \lambda)\hat{B}(b, \lambda)\rho(\lambda)d\lambda$$

where by definition $\int \rho(\lambda)d\lambda = 1$. The measuring instruments could also contain hidden variables which could affect the results, and therefore the results are averaged over. We denote the average of $\hat{A}$ and $\hat{B}$ over the instrument as A and B respectively, hence $|A| \leq 1$ and $|B| \leq 1$, so that

$$E(a, b) = \int A(a, \lambda)B(b, \lambda)\rho(\lambda)d\lambda.$$

Now, let a' and b' be alternate measurement settings, then

$$\begin{aligned} E(a, b) - E(a, b') &= \int \big(A(a, \lambda)B(b, \lambda) - A(a, \lambda)B(b', \lambda)\big)\, \rho(\lambda)d\lambda \\ &= \int A(a, \lambda)B(b, \lambda)[1 \pm A(a', \lambda)B(b', \lambda)]\rho(\lambda)d\lambda \\ &\quad - \int A(a, \lambda)B(b', \lambda)[1 \pm A(a', \lambda)B(b, \lambda)]\rho(\lambda)d\lambda. \end{aligned}$$

Using $|A| \leq 1, |B| \leq 1$ we note that

$$\begin{aligned} |E(a, b) - E(a, b')| &\leq \int [1 \pm A(a', \lambda)B(b', \lambda)]\rho(\lambda)d\lambda + \int [1 \pm A(a', \lambda)B(b, \lambda)]\rho(\lambda)d\lambda \\ &= 2\int \rho(\lambda)d\lambda \pm \int A(a', \lambda)B(b', \lambda)\rho(\lambda)d\lambda \pm \int A(a', \lambda)B(b, \lambda)\rho(\lambda)d\lambda \\ &= 2 \pm (E(a', b') + E(a', b)) \end{aligned}$$

Where the last step implies that $|E(a, b) - E(a, b')| \leq 2 - |E(a', b') + E(a', b)|$. Rearranging this yields $|E(a, b) - E(a, b')| + |E(a', b') + E(a', b)| \leq 2$, and by the triangle inequality we find $|E(a, b) - E(a, b') + E(a', b') + E(a', b)| \leq |E(a, b) - E(a, b')| + |E(a', b') + E(a', b)|$ which gives us the CHSH inequality,

$$|S| = |E(a, b) - E(a, b') + E(a', b') + E(a', b)| \leq 2. \tag{4}$$

4.3 Quantum Violation

We now deduce the quantum violation of the CHSH inequality following the presentation found in [15] and adding extra computational steps. Consider the singlet state $|\psi\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ and the observables

$$A = \sigma_z^A, \quad B = -\frac{1}{\sqrt{2}}(\sigma_z^B + \sigma_x^B), \quad A' = \sigma_x^A, \quad B' = \frac{1}{\sqrt{2}}(\sigma_z^B - \sigma_x^B).$$

Quantum mechanics predicts that

$$\begin{aligned}\langle A \otimes B\rangle_\psi &= \frac{1}{2\sqrt{2}}(\langle 01| - \langle 10|)(-\sigma_z^A \otimes \sigma_z^B - \sigma_z^A \otimes \sigma_x^B)(|01\rangle - |10\rangle) \\ &= \frac{1}{2\sqrt{2}}(\langle 01| - \langle 10|)(|01\rangle - |10\rangle - |00\rangle - |11\rangle) \\ &= \frac{1}{2\sqrt{2}}(\langle 01|01\rangle + \langle 10|10\rangle) = \frac{1}{\sqrt{2}},\end{aligned}$$

and similarly we can calculate the other expected values, which are

$$\langle A \otimes B\rangle_\psi = \frac{1}{\sqrt{2}}; \quad \langle A' \otimes B\rangle_\psi = \frac{1}{\sqrt{2}}; \quad \langle A' \otimes B'\rangle_\psi = \frac{1}{\sqrt{2}}; \quad \langle A \otimes B'\rangle_\psi = -\frac{1}{\sqrt{2}}.$$

Hence,

$$\langle A \otimes B\rangle_\psi + \langle A' \otimes B\rangle_\psi + \langle A' \otimes B'\rangle_\psi - \langle A \otimes B'\rangle_\psi = 2\sqrt{2} \not\leq 2, \tag{5}$$

which violates the CHSH inequality.

4.3.1 Experimental Confirmation

Although this result may seem counter-intuitive it is fundamental to our understanding of quantum theory and its description of the universe. We derived the CHSH inequality under the assumptions of locality and reality, that is the physical properties measured were predetermined and that Alice's measurements did not affect Bob's measurements. These two assumptions together are often referred to as local-realism. A number of experiments sought to prove this result. However, past experiments often had loopholes meaning that the experiment may not rule out some explanations or assumptions. Further informaton on loopholes can be found in [21, 19]. Clauser and Freedman held the first Bell test in 1972 [23] which was followed by Alain Aspect's famous optical test [24]. Notably, a paper by Handsteiner et al. [25] published in February 2017 detailed a loop-hole free experiment which measured polarisation-entangled photons from a distant astronomical source. The correlations between measurement outcomes were analysed using the CHSH inequality; the experiments found $S_{exp} = 2.425$ and $S_{exp} = 2.502$ which violates the bound derived under the assumption of local-realism .

5 Quantum Key Distribution

QKD uses quantum mechanics to establish secure lines of communication over public channels by generating a shared secret key between two parties. The first instance of associating a notion of secrecy to quantum theory was established by Wiesner in the 1970s [4] and published in 1983. Wiesner considered the idea of quantum money as a means to prevent counterfeiting. The first QKD protocol known as BB84 was established by Bennett and Brassard in 1984 [5] and is based on the disturbance of non-orthogonal states. In 1991, Artur Ekert proposed an entanglement-based QKD protocol known as Ekert 91 and subsequently many variants of both BB84 and Ekert 91 have been published. The No-Cloning Theorem was published in 1982 by Wootters and Zurek [26], was written in response to a paper incorrectly asserting that superluminal communication was possible [27].

Theorem 5.1 (No-Cloning Theorem)**.** *Let $|\psi\rangle \in \mathcal{H}_A$ be an unknown pure state and $|\alpha\rangle \in \mathcal{H}_B$ some known pure state. Suppose that we have some unitary operator $U \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ such that $U(|\psi\rangle \otimes |\alpha\rangle) = |\psi\rangle \otimes |\psi\rangle$ and $U(|\varphi\rangle \otimes |\alpha\rangle) = |\varphi\rangle \otimes |\varphi\rangle$. Then $|\psi\rangle = |\varphi\rangle$ or $|\psi\rangle$ and $|\varphi\rangle$ are orthogonal.*

Proof. By taking the inner product of $U(|\psi\rangle \otimes |\alpha\rangle) = |\psi\rangle \otimes |\psi\rangle$ and $U(|\varphi\rangle \otimes |\alpha\rangle) = |\varphi\rangle \otimes |\varphi\rangle$ the left hand side yields $(\langle\psi| \otimes \langle\alpha|)U^\dagger U(|\varphi\rangle \otimes |\alpha\rangle) = \langle\psi|\varphi\rangle \langle\alpha|\alpha\rangle = \langle\psi|\varphi\rangle$ and the right hand side yields $(\langle\psi| \otimes \langle\psi|) \otimes (|\varphi\rangle \otimes |\varphi\rangle) = \langle\psi|\varphi\rangle^2$. Hence, $\langle\psi|\varphi\rangle = \langle\psi|\varphi\rangle^2$ which implies that either $\langle\psi|\varphi\rangle = 1$ in which case $|\psi\rangle = |\varphi\rangle$ or $\langle\psi|\varphi\rangle = 0$ and $|\psi\rangle$ and $|\varphi\rangle$ are orthogonal as required. □

The No-Broadcasting theorem [28] generalises the No-Cloning theorem to mixed states. Below, the proposition that information gain implies disturbance means that any attempts made to eavesdrop will introduce some disturbance to a signal. In 5.2 an example of an eavesdropping attack is given where Eve's strategy is to hide the disturbance she introduces behind environmental noise.

Proposition 5.1 (Information Gain Implies Disturbance)**.** *In any attempt to distinguish between two non-orthogonal quantum states, information gain is only possible at the expense of introducing disturbance to the signal.*

Proof. Here we follow [15]. Suppose that $|\psi\rangle$ and $|\varphi\rangle$ are non-orthogonal quantum states. Without loss of generality Eve attempts to obtain information about the states by unitarily interacting the state $|\psi\rangle$ and $|\varphi\rangle$ with an ancilla in state $|u\rangle$. If this process does not introduce any disturbance then $|\psi\rangle |u\rangle \mapsto |\psi\rangle |v\rangle$ and $|\varphi\rangle |u\rangle \mapsto |\varphi\rangle |v'\rangle$. Since inner products are preserved under unitary transformations we have that $\langle v|v'\rangle \langle u|u\rangle \langle\psi|\varphi\rangle = \langle u|u\rangle \langle\psi|\varphi\rangle$ and thus $\langle v|v'\rangle = \langle u|u\rangle = 1$ so Eve cannot distinguish between the states $|\psi\rangle$ and $|\varphi\rangle$. □

5.1 Ekert 91

In 1991 Artur Ekert devised a way of using quantum entanglement as a resource for quantum key distribution, in most literature this is referred to as Ekert 91 or the EPR Protocol [6]. The security of the protocol depends on quantum mechanics being a complete theory, whereby there are no hidden variables and quantum mechanics provides the maximal amount of information about the system.

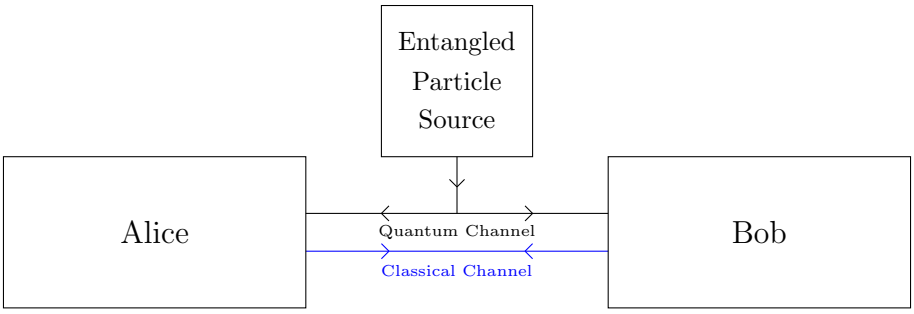

Figure 1: Ekert 91 Protocol

The scheme Ekert presented uses the CHSH inequality detailed in section 4 to test for eavesdropping [29]. In this section we will explain the Ekert 91 protocol and how the CHSH inequality can be used to detect eavesdropping. Due to the probabilistic nature of quantum mechanics, measurements of a singlet state are perfectly random, that is the probability of measuring $|0\rangle$ or $|1\rangle$ is one half. This means that a perfectly random binary key can be generated using Ekert 91 for use in the provably secure one time pad.

The Ekert 91 Protocol

1. A source emits pairs of spin-$\frac{1}{2}$ particles, which are maximally entangled and sends one to Alice and one to Bob along the y axis over the quantum channel.

2. Alice and Bob randomly choose one of three unit vectors to perform a measurement on the spin components of each particle. The unit vectors a_i and b_j lie in the $x-z$ plane perpendicular to the direction the particle arrives from, with angles $\theta_1^a = 0,\ \theta_2^a = \frac{1}{4}\pi,\ \theta_3^a = \frac{1}{2}\pi,\ \theta_1^b = \frac{1}{4}\pi,$ $\theta_2^b = \frac{1}{2}\pi,\ \ \theta_3^b = \frac{3}{4}\pi$. Here a and b are used to label Alice and Bob's analysers respectively.

3. Alice and Bob publicly announce over a classical channel the orientations of the analysers used to perform each measurement. They do not announce the results of the measurements. Measurements are divided into two separate groups. Measurements in group 1 were performed using different analyser orientations. Measurements in group 2 were performed using the same analyser orientation. Measurements where a particle was not detected are discarded.

4. Alice and Bob publicly compare the measurement results over the classical channel from the first group only. Using this they establish a value for

$$S = E(a_1, b_1) - E(a_1, b_3) + E(a_3, b_1) + E(a_3, b_3);$$

if the particles were not disturbed by a potential eavesdropper then $S = -2\sqrt{2}$ and the key is assumed safe to use.

After this there are two important steps known as key reconciliation and privacy amplification. In the remaining group of particles their may be errors introduced through the measurement apparatus or environmental noise. In the key reconciliation stage Alice and Bob can use error correction

techniques to remove errors within the key and establish a shared *raw key*. An introduction to these techniques can be found in [15]. Quantum privacy amplification can then be applied to reduce the information Eve knows about the key, thus this is considered to amplify Alice and Bob's privacy.

5.1.1 Ekert 91 in Practice

In section 4 we concisely demonstrated a violation of the CHSH inequality. In this section we will calculate violations of the CHSH inequality within the setting of Alice and Bob attempting to establish a secret key using the Ekert 91 protocol. These calculations will be essential in understanding the arguments constructed in sections 5.2 and 5.3. We now define the *correlation coefficient* which can be used in experimental settings.

Definition 5.1 (Correlation Coefficient)**.** *The correlation coefficient of the measurements performed by Alice and Bob in the directions* $\mathbf{a_i}$ *and* $\mathbf{b}_j$ *respectively is*

$$E(\mathbf{a_i}, \mathbf{b_j}) = P_{++}(\mathbf{a_i}, \mathbf{b_j}) + P_{--}(\mathbf{a_i}, \mathbf{b_j}) - P_{+-}(\mathbf{a_i}, \mathbf{b_j}) - P_{-+}(\mathbf{a_i}, \mathbf{b_j}).$$

Here $P_{\pm\pm}$ denotes the probability that the result ± 1 has been obtained along $\mathbf{a}_i$ and $\mathbf{b}_j$, respectively. It is clear that the probability of Alice or Bob measuring either ± 1 is $P_{\pm}(\mathbf{a_i}) = P_{\pm}(\mathbf{b_j}) = \frac{1}{2}$. We will now compute the the correlation coefficient, which is the quantum expectation

$$E(a_i, b_j) = \langle \mathbf{a_i} \cdot \boldsymbol{\sigma} \otimes \mathbf{b_j} \cdot \boldsymbol{\sigma} \rangle = -\mathbf{a_i} \cdot \mathbf{b_j}.$$

We now suppress the subscripts. First we expand the operator $\mathbf{a} \cdot \boldsymbol{\sigma} \otimes \mathbf{b} \cdot \boldsymbol{\sigma}$

$$\begin{aligned} W := (\mathbf{a} \cdot \boldsymbol{\sigma}) \otimes (\mathbf{b} \cdot \boldsymbol{\sigma}) = {} & a_1 b_1 \sigma_x \otimes \sigma_x + a_1 b_2 \sigma_x \otimes \sigma_y + a_1 b_3 \sigma_x \otimes \sigma_z \\ & + a_2 b_1 \sigma_y \otimes \sigma_x + a_2 b_2 \sigma_y \otimes \sigma_y + a_2 b_3 \sigma_y \otimes \sigma_z \\ & + a_3 b_1 \sigma_z \otimes \sigma_x + a_3 b_2 \sigma_z \otimes \sigma_y + a_3 b_3 \sigma_z \otimes \sigma_z. \end{aligned} \tag{6}$$

For use in the following calculation recall that

$$\sigma_x |0\rangle = |1\rangle, \sigma_x |1\rangle = |0\rangle, \ \sigma_y |0\rangle = i |1\rangle, \ \sigma_y |1\rangle = -i |0\rangle, \ \sigma_z |0\rangle = |0\rangle \text{ and } \sigma_z |1\rangle = -|1\rangle,$$

and thus we compute $\sqrt{2} W |\psi^-\rangle$ as follows:

$$\begin{aligned} \sqrt{2} W |\psi^-\rangle = {} & a_1 b_1 |10\rangle - i a_1 b_2 |10\rangle - a_1 b_3 |11\rangle - a_1 b_1 |01\rangle - i a_1 b_2 |01\rangle - a_1 b_3 |00\rangle \\ & + i a_2 b_1 |10\rangle + a_2 b_2 |10\rangle - i a_2 b_3 |11\rangle + i a_2 b_1 |01\rangle - a_2 b_2 |01\rangle + i a_2 b_3 |00\rangle \\ & + a_3 b_1 |00\rangle - i a_3 b_2 |00\rangle - a_3 b_3 |01\rangle + a_3 b_1 |11\rangle + i a_3 b_2 |11\rangle + a_3 b_3 |10\rangle. \end{aligned} \tag{7}$$

Note that the first three terms of each line correspond to the operators acting upon $|01\rangle$ and next three terms correspond to $|10\rangle$. Now we compute $\langle \psi^- | W | \psi^- \rangle$. Note that $\langle ij|kl\rangle = \langle i|k\rangle \langle j|l\rangle = \delta_{ik}\delta_{jl}$, and states of the form $\langle ij|ii\rangle = 0$, thus we can ignore kets of the form $|ii\rangle$.

$$\begin{aligned} \langle \psi^- | W | \psi^- \rangle = {} & \frac{1}{2}(-a_1 b_1 \langle 10|10\rangle + i a_1 b_2 \langle 10|10\rangle - a_1 b_1 \langle 01|01\rangle - i a_1 b_2 \langle 01|01\rangle \\ & - i a_2 b_1 \langle 10|10\rangle - a_2 b_2 \langle 10|10\rangle + i a_2 b_1 \langle 01|01\rangle - a_2 b_2 \langle 01|01\rangle \\ & - a_3 b_3 \langle 01|01\rangle - a_3 b_3 \langle 10|10\rangle) = -\frac{1}{2} 2 (a_1 b_1 + a_2 b_2 + a_3 b_3) \\ = {} & -\mathbf{a} \cdot \mathbf{b} = -\cos\theta_{ab} \end{aligned} \tag{8}$$

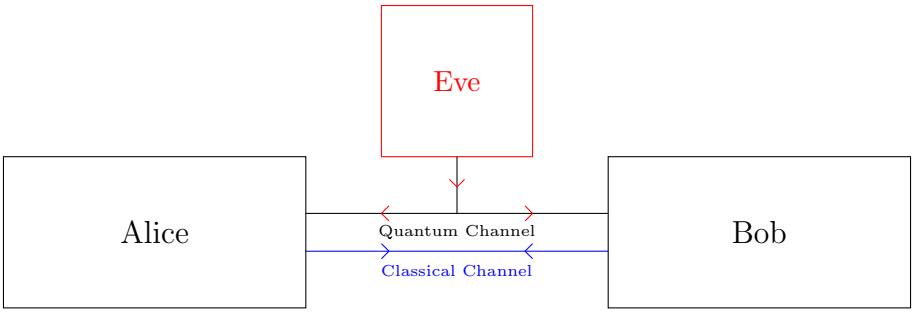

Figure 2: Eavesdropping on the Ekert 91 Protocol

here**a** and **b** are unit vectors, θ_{ab} denotes the angle between **a** and **b**. Using this we can calculate $E(a_i, b_j)$ for various orientations.

If the angle between Alice and Bob's analysers is zero there will be a complete anti-correlation between the measurement results since $E(a_2, b_1) = E(a_3, b_2) = -\cos 0 = -1$. For the other orientations we have $\theta_{a_1 b_1} = \theta_{a_3 b_1} = \theta_{a_3 b_3'} = \frac{\pi}{4}$, $\theta_{a_1 b_3} = \frac{3\pi}{4}$. Therefore $\cos\theta_{a_1 b_1} = \cos\frac{\pi}{4} = \frac{\sqrt{2}}{2}$ and $\cos\theta_{a_1 b_3} = -\cos\frac{3\pi}{4} = -\frac{\sqrt{2}}{2}$. This yields the important result

$$S = E(a_1, b_1) - E(a_1, b_3) + E(a_3, b_1) + E(a_3, b_3) = -4\frac{2}{\sqrt{2}} = -2\sqrt{2}. \tag{9}$$

From equation (9) we can clearly see that the first group of measurements, performed with different orientations, maximally violate the Bell inequality and therefore can be used to test for eavesdropping [21].

5.2 Eavesdropping

Just as classical signals can be disturbed in transmission by the environment, so can quantum communications. If Alice and Bob are trying to establish a secure key and are checking for CHSH violations, particles will naturally encounter environmental noise that will affect some measurements. When Eve eavesdrops by 5.1 she will introduce a disturbance; one strategy Eve can employ is to try and make this disturbance appear to be environmental noise. To do this, Eve needs to establish a method of estimating Alice and Bob's measurements whilst minimising the amount of disturbance she introduces in an effort to imitate environmental noise. We will follow and add detail to the outlined eavesdropping attack described by Ekert in [21]. As shown schematically in figure 2, Eve controls the particle source and therefore can prepare the particles to her advantage. We will consider the case when environmental noise is said to be symmetric in the $x-z$ plane [30], and the correlation coefficient takes the form

$$E(\mathbf{a}, \mathbf{b}) = \langle(\mathbf{a}\cdot\boldsymbol{\sigma}) \otimes (\mathbf{b}\cdot\boldsymbol{\sigma})\rangle = -\eta\ \mathbf{a}\cdot\mathbf{b}, \tag{10}$$

where $\eta \in [0, 1]$ [21]. Here **a** and **b** are unit vectors in the $x-z$ plane as stated in 5.1. It can be shown that the correlation coefficient can only take this form if the reduced density matrix after

tracing out Eve's subsystem is

$$\rho_{AB} = A\left|\psi^-\right\rangle\left\langle\psi^-\right| + B\left|\phi^+\right\rangle\left\langle\phi^+\right| + \frac{C}{4}\mathbb{I}. \tag{11}$$

This form follows from the fact that that the Bell states are rotationally invariant in the $x-z$ plane, clearly the identity matrix is rotationally invariant which ensures that the noise is symmetrical in the $x-z$ plane. To see that $|\psi^-\rangle$ is rotationally invariant (up to a global phase factor), and by extension $|\psi^-\rangle\langle\psi^-|$) is rotationally invariant, it suffices to show that $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|vv^\perp\rangle - |v^\perp v\rangle)$, where $|v^\perp\rangle$ denotes a state orthogonal to $|v\rangle$. Let $|v\rangle = \alpha|0\rangle + \beta|1\rangle$ and choose $|v^\perp\rangle = \alpha^*|1\rangle - \beta^*|0\rangle$. These states are orthogonal since $\langle v^\perp|v\rangle = 0$. Then $(|vv^\perp\rangle - |v^\perp v\rangle) = (|\alpha|^2+|\beta|^2)|01\rangle - (|\alpha|^2+|\beta|^2)|10\rangle = |01\rangle - |10\rangle$ as required.

Here ρ_{AB} is written in the Bell basis. Using calculations from section 5.1.1 and the fact that Pauli spin matrices have trace 0 the expected value of the reduced density operator is

$$\begin{aligned}
\mathrm{Tr}\left[\rho_{AB}(\mathbf{a}\cdot\boldsymbol{\sigma})\otimes(\mathbf{a}\cdot\boldsymbol{\sigma})\right] &= \mathrm{Tr}\left[\left(A\left|\psi^-\right\rangle\left\langle\psi^-\right| + B\left|\phi^+\right\rangle\left\langle\phi^+\right| + \frac{C}{4}\mathcal{I}\right)(\mathbf{a}\cdot\boldsymbol{\sigma})\otimes(\mathbf{a}\cdot\boldsymbol{\sigma})\right] \\
&= A\left\langle\psi^-\right| W\left|\psi^-\right\rangle + B\left\langle\phi^+\right| W\left|\phi^+\right\rangle + \mathrm{Tr}\frac{C}{4}(\mathcal{I}_2\otimes\mathcal{I}_2)W \\
&= -A(\mathbf{a}\cdot\mathbf{b}) + B(\mathbf{a}\cdot\mathbf{b}) = -(A-B)(\mathbf{a}\cdot\mathbf{b}).
\end{aligned}$$

Hence the coefficients A, B of the reduced density operator ρ_{AB} are equal to the disturbance, that is

$$A - B = \eta. \tag{12}$$

Without loss of generality Eve can prepare a pure state to eavesdrop on Alice and Bob of the form $|\varphi\rangle = \sum_{\alpha,\beta\in\{0,1\}}|\alpha\rangle|\beta\rangle|\hat{E}_{\alpha,\beta}\rangle$ [31]. Here Eve has prepared the two particles with an ancilla E and the states $|\alpha\rangle$ and $|\beta\rangle$ refer to Alice and Bob's states respectively. Eve's states are not necessarily normalised or orthogonal and the only necessary condition on Eve's state is the overall normalisation of $|\varphi\rangle$ such that $\sum_{\alpha,\beta\in\{0,1\}}\langle\hat{E}_{\alpha,\beta}|\hat{E}_{\alpha,\beta}\rangle = 1$ [32]. Thus, Eve prepares the state

$$|\zeta\rangle_{ABE} = \frac{\sqrt{F}}{\sqrt{2}}(|01\rangle|E_{01}\rangle + |10\rangle|E_{10}\rangle) + \frac{\sqrt{D}}{\sqrt{2}}(|00\rangle|E_{00}\rangle + |11\rangle|E_{11}\rangle), \tag{13}$$

where $|E_{\alpha,\beta}\rangle$ is the normalised version of the state $|\hat{E}_{\alpha,\beta}\rangle$, with density operator

$$\begin{aligned}
\rho_{ABE} = &\left[\frac{\sqrt{F}}{\sqrt{2}}(|01\rangle|E_{01}\rangle + |10\rangle|E_{10}\rangle) + \frac{\sqrt{D}}{\sqrt{2}}(|00\rangle|E_{00}\rangle + |11\rangle|E_{11}\rangle)\right] \\
&\left[\frac{\sqrt{F}}{\sqrt{2}}(\langle 01|\langle E_{01}| + \langle 10|\langle E_{10}|) + \frac{\sqrt{D}}{\sqrt{2}}(\langle 00|\langle E_{00}| + \langle 11|\langle E_{11}|)\right]
\end{aligned} \tag{14}$$

Tracing over the ancilla we find the reduced density matrix $\mathrm{tr}_E[\rho_{ABE}] = \rho_{AB}$ if

$$F = \frac{1}{2}(1 + A - B), \quad D = \frac{1}{2}(1 - A + B)$$

and

$$\langle E_{01}|E_{10}\rangle = \frac{A}{F} = \cos\alpha, \quad \langle E_{00}|E_{11}\rangle = \frac{B}{D} = \cos\beta, \tag{15}$$

this is a convenient parameterisation as seen in [33] for the overlap of these non-orthogonal states, the remaining inner products are zero. Thus, $\{|E_{01}\rangle, |E_{10}\rangle\}$ and $\{|E_{00}\rangle, |E_{11}\rangle\}$ are orthogonal subspaces. Note that the inner product is the same regardless of what basis $|E_{kl}\rangle$ is written in; consider the unitary transform $U : |E_{kl}\rangle \mapsto U|E_{kl}\rangle$ then $(|E_{kl}\rangle)^\dagger \mapsto (U|E_{kl}\rangle)^\dagger = \langle E_{kl}|U^\dagger$ and $\langle E_{ij}|U^\dagger U|E_{kl}\rangle = \langle E_{ij}|E_{kl}\rangle$.

Eve observes Alice and Bob compare the orientations in which they took their measurements. Alice and Bob communicate over the public channel and as stated previously compare the measurement results of the first group to determine a value for S. Alice and Bob will have a predetermined minimal value of S that assures them it is safe to continue with the protocol, this minimal value will account for environmental noise. Assuming Eve has successfully imitated environmental noise and Alice and Bob have achieved a sufficiently high value for S the protocol continues. Alice, Bob and Eve now discard the first group.

In the group that remains, are the measurements that were performed with different orientations. Eve knows the orientation in which the measurements were taken from observing the public channel and now attempts to identify the state of her ancilla. We have seen in example 3.4 that non-orthogonal states cannot be reliably distinguished and since the states of the ancilla are not necessarily orthogonal there is some probability of identifying the state incorrectly. The Helstrom Bound [34] states that the minimal probability of error for an optimal measurement that can discriminate between two non-orthogonal pure states $|\psi_i\rangle$ and $|\psi_j\rangle$ with a priori-probabilities μ_i and μ_j is

$$\frac{1}{2}\left(1 - \sqrt{1 - 4\mu_1\mu_2|\langle\psi_i|\psi_j\rangle|^2}\right)$$

and the probability of correctly identifying the state is therefore

$$1 - \frac{1}{2}\left(1 - \sqrt{1 - 4\mu_1\mu_2|\langle\psi_i|\psi_j\rangle|^2}\right) = \frac{1}{2}\left(1 + \sqrt{1 - 4\mu_1\mu_2|\langle\psi_i|\psi_j\rangle|^2}\right).$$

We will not concern ourselves with how to deduce such a measurement, however this can be found in [34]. In our case $\mu_1 = \mu_2 = \frac{1}{2}$ and therefore the probability of incorrectly distinguishing the state is

$$\frac{1}{2}\left(1 - \sqrt{1 - |\langle E_{ij}|E_{kl}\rangle|^2}\right).$$

Eve can eavesdrop on Alice and Bob's measurements with an error rate of

$$\begin{aligned} Q_E &= \frac{F}{2}\left(1 - \sqrt{1 - |\langle E_{01}|E_{10}\rangle|^2}\right) + \frac{D}{2}\left(1 - \sqrt{1 - |\langle E_{00}|E_{11}\rangle|^2}\right) \\ &= \frac{F}{2}(1 - \sqrt{1 - \cos^2\alpha}) + \frac{D}{2}(1 - \sqrt{1 - \cos^2\beta}) \\ &= \frac{F}{2}(1 - \sin\alpha) + \frac{D}{2}(1 - \sin\beta). \end{aligned} \tag{16}$$

Rearranging equation (15) we find the disturbance $\eta = A - B = F\cos\alpha - D\cos\beta$; this error can be minimised by choosing $\cos\alpha = -\cos\beta$ and thus the disturbance is

$$\eta = A - B = (F + D)\cos\alpha = \cos\alpha.$$

Eve's error rate is therefore

$$Q_E = \frac{F + D}{2}(1 - \sqrt{1 - \cos^2\alpha}) = \frac{1}{2}(1 - \sin\alpha),$$

and the error rate introduced to the key by the disturbance $\eta = \cos\alpha$ is then $Q_{AB} = \frac{1}{2}(1 - \cos\alpha)$ which is equal to Eve's error when $\alpha = \frac{\pi}{4}$. When the error rates are equal $E(\mathbf{a}, \mathbf{b}) = -\eta\ \mathbf{a}\cdot\mathbf{b} = \frac{1}{\sqrt{2}}$, and remarkably

$$|S| = \frac{1}{\sqrt{2}}|E(a_1, b_1) - E(a_1, b_3) + E(a_3, b_1) + E(a_3, b_3)| = 2.$$

Here the classical bound of the CHSH inequality marks the crossing point of the two error rates. One way of interpreting this, as noted by Ekert [21], is the maximal acceptable error rate in which Alice and Bob can establish a secure key using some error correction techniques and with no additional public communication. Or, if $|S| > 2$ then Alice and Bob can proceed in communications and have established a secure key. In the following section we will consider whether secure keys can be established for $|S| < 2$.

5.3 Quantum Privacy Amplification

Quantum privacy amplification is a process that allows Alice and Bob to reduce the amount of information Eve may have about their secret key. Eve could have gathered this information by eavesdropping on the quantum channel or over the public channel during the information reconciliation stage when Alice and Bob corrected errors in the key. Privacy amplification uses the error corrected key to produce a shorter key that Eve has less information about. This can be achieved using the process of entanglement purification. Bennett et al. proposed a purification scheme in [35] as a method of achieving faithful transmission or removing noise from a transmission. The scheme achieves this by converting mixed entangled states to near perfect entangled states, that is states close in form to Bell states. We will present the basic idea of the quantum privacy amplification method published by Ekert et al. [10]. A complete technical description of purification-based privacy amplification is beyond the scope of this report, however, we will discuss how applying QPA can result in the interesting result that Alice and Bob can generate a secure key using Ekert 91 for non-violations of the CHSH inequality.

Suppose that Eve prepares two qubit pairs and sends one qubit from each pair to Alice and Bob. Let the density operators of the two pairs be ρ and ρ'. Next, Alice performs the unitary operation U_A and Bob performs U_B on the two qubits.

$$U_A = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}, \quad U_B = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}, \quad U_A U_B = \mathbb{I}$$

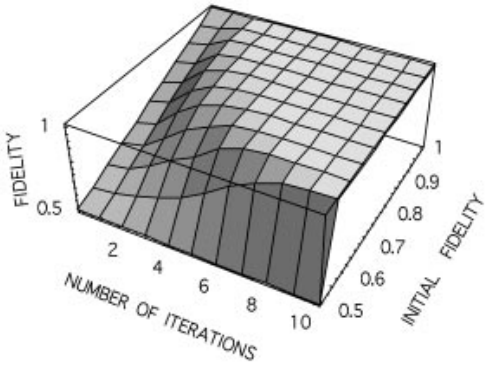

Figure 3: Fidelity using QPA. Graph source: [10].

Alice and Bob then perform two quantum CNOT operations,

$$|x\rangle |y\rangle \rightarrow |x\rangle |x \oplus y\rangle , \quad (x, y) \in \{0, 1\}.$$

For example, $|1\rangle |0\rangle \rightarrow |1\rangle |1 \oplus 0\rangle = |1\rangle |1\rangle$ and $|1\rangle |1\rangle \rightarrow |1\rangle |1 \oplus 1\rangle = |1\rangle |0\rangle$, here $\oplus$ is the same as addition modulo 2. Both Alice and Bob then measure the z-components of the target's spin. If both of the outcomes coincide, for both are $|1\rangle$, as in the example above, then the control pair are kept for the next round and the target is discarded. If the outcomes do not coincide the pairs are discarded.

Let the density operator ρ be the state of one pair of particles. We can express the state in the Bell basis with diagonal elements A, B, C, D such that

$$A = \langle \phi^+ | \rho | \phi^+ \rangle , \quad B = \langle \psi^- | \rho | \psi^- \rangle$$

$$C = \langle \psi^+ | \rho | \psi^+ \rangle , \quad D = \langle \phi^- | \rho | \phi^- \rangle .$$

A is said to be the *fidelity*, and can be considered a measure of how close a qubit is to being in the state $|\phi^+\rangle$, we want A to be close to one and the other diagonal elements close to zero. The process is carried out on an ensemble of such pairs and the average diagonal entries of the surviving pairs are calculated. Suppose that the two pairs of two particles are in the state $\rho \otimes \rho$. If measurements of the target pair coincide, the controls are kept with density operator $\tilde{\rho}$. Then $\tilde{\rho}$ has diagonal entries $\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}$, such that

$$\tilde{A} = \frac{A^2 + B^2}{N}, \quad \tilde{B} = \frac{2CD}{N}, \quad \tilde{C} = \frac{C^2 + D^2}{N}, \tilde{D} = \frac{2AB}{N}$$

and $N = (A + B)^2 + (C + D)^2$ is the probability of Alice and Bob's measurements coinciding. This process would obviously be undertaken with a large sample of pairs. The average fidelity is then found and the process is repeated for surviving pairs. The application of this process is shown in figure 3, it is clear that a large number of pairs would be required to establish a secure key as many will be discarded. The QPA procedure reduces the information Eve has access too by purifying

mixed states to pure states. A complete description of this process can be found in [10]. Here we have presented a narrow overview of purification based QPA, we will show an application of such a process in the following section.

5.3.1 Establishing Secure Keys with QPA

With this process in mind and following [21], recall the reduced density operator in equation (11) and set the coefficients

$$A = \frac{1}{2}(\cos\alpha)(1+\cos\alpha), \quad B = -\frac{1}{2}(\cos\alpha)(1-\cos\alpha),$$

and $C = \frac{1}{4}\sin^2\alpha$ to ensure the state is normalised. Thus we have the state

$$\rho(\alpha) = \frac{1}{2}(\cos\alpha)(1+\cos\alpha)\left|\psi^-\right\rangle\left\langle\psi^-\right| - \frac{1}{2}(\cos\alpha)(1-\cos\alpha)\left|\Phi^+\right\rangle\left\langle\Phi^+\right| + \frac{1}{4}(\sin^2\alpha)\mathbb{I},$$

where ρ is dependent on α. By equation (10) and (12) we have $E(a,b) = -(A-B)\mathbf{a}\cdot\mathbf{b} = -(\cos\alpha)\mathbf{a}\cdot\mathbf{b}$ and therefore

$$|S| = |(\cos\alpha)2\sqrt{2}|.$$

Using the PPT criterion 3.2 we can find the minimal value of $\cos\alpha$ such that $\rho(\alpha)$ is an entangled state and apply quantum privacy amplification states of this form. The partial transpose of $\rho(\alpha)$ is

$$\begin{aligned}\rho^{T_B}(\alpha) = {} & (B+C)\left|0\right\rangle\left\langle 0\right| \otimes \left|0\right\rangle\left\langle 0\right| + (B+C)\left|1\right\rangle\left\langle 1\right| \otimes \left|1\right\rangle\left\langle 1\right| + (A+C)\left|0\right\rangle\left\langle 0\right| \otimes \left|1\right\rangle\left\langle 1\right| \\ & + (A+C)\left|1\right\rangle\left\langle 1\right| \otimes \left|0\right\rangle\left\langle 0\right| - A\left|0\right\rangle\left\langle 1\right| \otimes \left(\left|1\right\rangle\left\langle 0\right|\right)^T - A\left|1\right\rangle\left\langle 0\right| \otimes \left(\left|0\right\rangle\left\langle 1\right|\right)^T \\ & + B(\left|0\right\rangle\left\langle 1\right| \otimes \left(\left|0\right\rangle\left\langle 1\right|\right)^T + B\left|1\right\rangle\left\langle 0\right| \otimes \left(\left|1\right\rangle\left\langle 0\right|\right)^T,\end{aligned}$$

or equivalently, in matrix notation ρ and ρ^{T_B} are

$$\rho = \frac{1}{2}\begin{pmatrix} B+C & 0 & 0 & B \\ 0 & A+C & -A & 0 \\ 0 & -A & A+C & 0 \\ B & 0 & 0 & B+C \end{pmatrix}, \quad \rho^{T_B} = \frac{1}{2}\begin{pmatrix} B+C & 0 & 0 & -A \\ 0 & A+C & B & 0 \\ 0 & B & A+C & 0 \\ -A & 0 & 0 & B+C \end{pmatrix}.$$

Calculating the eigenvalues of $\rho(\alpha)$, $(\det(\rho^{T_B} - \lambda\mathbb{I}) = 0)$, by the PPT criterion3.2 we require that at least one eigenvalue is negative so that $\rho(\alpha)$ is an entangled state. The smallest eigenvalue is $\lambda = \sqrt{2} - 1 - \cos\alpha$ and therefore $\rho(\alpha)$ is entangled if $\cos\alpha > \sqrt{2} - 1$. Applying the quantum privacy amplification outlined in section 5.3 to states of this form will result in states converging to a Bell state. This implies that for a sufficient amount of states, Alice and Bob could establish a secure key with a non-violation of CHSH inequality, that is

$$|S| > (\sqrt{2}-1)2\sqrt{2} = 4 - 2\sqrt{2}.$$

6 Conclusion

In this project, we have reviewed the trouble with modern day public key cryptography being built upon computationally hard problems and reviewed how quantum computers and the implementation of Shor's algorithm will exacerbate this problem. This led us to consider quantum key distribution as a solution to this problem, and in particular, an entanglement-based quantum key distribution protocol known as Ekert 91. This report aimed to introduce an undergraduate reader into entanglement-based quantum key distribution, which is not as accessible and well documented as BB84. Thus we introduced the relevant background mathematics including quantum entanglement, state discrimination and the positive partial transpose separability test. This established the mathematical foundations required to understand the following sections.

Next, the EPR paradox and Bohm's simplification of the paradox were introduced to highlight the fundamental nature of entanglement and its implications in our understanding of the world around us. The generalised Bell inequality known as the CHSH inequality was then derived, which places bounds on the statistical correlations of measurement outcomes in theories which assume locality and reality. It was then explicitly shown that correlations due to quantum entanglement violate the CHSH inequality and that the 2017 loop-hole free Handsteiner et al. experiment [25] supports this claim.

Following this, quantum key distribution was introduced along with two important results, the no-cloning theorem and that information gain implies disturbance. The Ekert 91 protocol was then explained showing how the CHSH inequality can be used to test for eavesdropping. It was noted how the inherent randomness of quantum mechanics can be used to generate truly random keys for use in the provably secure one time pad.

By example we illustrated an eavesdropping attack that imitates environmental noise. This presented Eve with the problem of discriminating between non-orthogonal states and how an optimal measurement can be chosen to minimise her probability of error as stated by the Helstrom bound. Lastly we illustrated how quantum privacy amplification can be used to generate a shared secret key when presented with eavesdropping attempts.

6.1 Outlook

The first implementation of the Ekert 91 protocol was performed by using polarised-entangled photons. There have since been efforts to commercially implement QKD for use by government agencies. Ekert 91 is an example of a discrete variable QKD protocol. The future of QKD lies with continuous variable protocols [36] as demonstrated by Jouguet et al. [37] in their recent experiments. The experiments showed that quantum devices will be compatible with our current communication infrastructure. The experiment successfully implemented continuous variable quantum key distribution over 80 kilometres through optical fibres and therefore we can expect to see QKD securing our communications in the future.

References

[1] Delfs, H. and Knebl, H.; *Introduction to cryptography Vol. 3*; Springer-Verlag Berlin Heidelberg; 2015.

[2] Kaliski, B.; *TWIRL and RSA key size*; 2003.

[3] Shor, P. W.; *Algorithms for quantum computation: Discrete logarithms and factoring*; Foundations of Computer Science, 1994 Proceedings, 35th Annual Symposium on Foundations of Computer Science: 124-134; 1994.

[4] Wiesner, S.; *Conjugate coding*; ACM Sigact News 15(1): 78-88; 1983.

[5] Bennett, C. H. and Brassard, G.; *Quantum cryptography: Public key distribution and coin tossing*; Proceedings of IEEE International Conference on Computers, Systems and Signal Processing 175: 8; 1984.

[6] Ekert, A. K.; *Quantum cryptography based on Bell's Theorem*; Physical Review Letters 67(6): 661-663; 1991.

[7] Einstein, A., Podolsky, B. and Rosen, N.; *Can quantum-mechanical description of physical reality be considered complete?*; Physical Review 47: 777; 1935.

[8] Bell, J. S.; *On the Einstein Podolsky Rosen Paradox*; Physics 1(3): 195-200; 1964.

[9] Clauser, J. F., Horne, M. A., Shimony, A. and Holt, R. A.; *Proposed experiment to test local hidden-variable theories*; Physical Review Letters 23: 880; 1969.

[10] Deutsch, D., Ekert, A., Jozsa, R., Macchiavello, C., Popescu, S. and Sanpera, A.; *Quantum privacy amplification and the security of quantum cryptography over noisy channels*; Physical Review Letters 77(13): 2818; 1996.

[11] Stanoyevitch, A.; *Introduction to cryptography with mathematical foundations and computer implementations, 1st ed.*; Chapman and Hall/CRC; 2013.

[12] Biham, E.; *Advances in cryptology - EUROCRYPT 2003, 1st ed.*; Berlin: Springer; 2003.

[13] Shannon, C. E.; *Communication theory of secrecy systems*; Bell Labs Technical Journal 28(4): 656-715; 1949.

[14] Heinosaari, T. and Ziman, M.; *The mathematical language of quantum theory from uncertainty to entanglement*; Cambridge University Press; 2012.

[15] Nielson, M. A. and Chuang, I. L.; *Quantum computation and quantum information*; Cambridge Univ. Press, Cambridge; 2000.

[16] Satish, S. and Vasudeva, H.; *Metric spaces*; Springer, London; 2006.

[17] http://www.unm.edu/~roy/usd/usd_review.pdf; (Accessed: 03/04/2017)

[18] Peres, A.; *Separability criterion for density matrices*; Physical Review Letters 77(8): 1413; 1996.

[19] Gennaro, A.; *Foundations and interpretation of quantum mechanics: In the light of a critical-historical analysis of the problems and of a synthesis of the results*; World Scientific; 2001.

[20] Bohm, D.; *Quantum theory*; Prentice-Hall Englewood Cliffs; 1951.

[21] Bertlmann, R. A. and Zeilinger, A., editors; *Quantum (Un)speakables: From Bell to Quantum Information*; Springer Science and Business Media; 2013.

[22] Bell, J.; *Introduction to the hidden-variable question*; Foundations of Quantum Mechanics: Proceedings of the International School of Physics: 171-81; 1971.

[23] Freeman, S. J. and Clauser, J.; *Experimental tests of local hidden-variable theories*; Physical Review Letters 28: 938-941; 1972.

[24] Aspect, A., Grangier, P. and Roger, G.; *Experimental tests of realistic local theories via Bell's Theorem*; Physical Reviews Letters 47(7): 460–3; 1981.

[25] Handsteiner, J., Friedman, A. S., Rauch, D., Gallicchio, J., Liu, B., Hosp, H., Kofler, J., Bricher, D., Fink, M., Leung, C. and Mark, A.; *Cosmic Bell test: Measurement settings from Milky Way stars*; Physical Review Letters 118(6): 060401; 2017.

[26] Wootters, W. K. and Zurek, W. H.; *A single quantum cannot be cloned*; Nature 299(5886): 802-803; 1982.

[27] Peres, A; *How the no-cloning theorem got its name*; Fortschritte der Physik 51(4-5): 458-461; 2003.

[28] Barnum, H., Caves, C. M., Fuchs, C. A., Jozsa, R. and Schumacher, B.; *Noncommuting mixed states cannot be broadcast*; Physical Review Letters 76(15): 2818; 1996.

[29] Bouwmeester, D., Ekert, A. and Zeilinger, A.; *The physics of quantum information*; Springer, Berlin; 2000.

[30] Durt, T., Kaszlikowski, D., Chen, J. L. and Kwek, L. C.; *Security of quantum key distributions with entangled qudits*; Physical Review A 69(3): 032313; 2004.

[31] Uhlmann, A.; *The transition probability in the state space of a *-algebra*; Reports on Mathematical Physics 9(2): 273; 1976.

[32] Inamori, H., Rallan, L. and Vedral, V.; *Security of EPR-based quantum cryptography against incoherent symmetric attacks*; Journal of Physics A: Mathematical and General 34(35): 6913; 2001.

[33] Cirac, J. I. and Gisin, N.; *Coherent eavesdropping strategies for the four state quantum cryptography protocol*; Physics Letters A 229(1): 1-7; 1997.

[34] Helstrom, C. W.; *Quantum detection and estimation theory*; Academic Press; 1976.

[35] Bennett, C. H., Brassard, G., Popescu, S., Schumacher, B., Smolin, J. A. and Wootters, W. K.; *Purification of noisy entanglement and faithful teleportation via noisy channels*; Physical Review Letters 76(5): 722; 1996.

[36] Weedbrook, C. et al.; *Gaussian quantum information*; Reviews of Modern Physics 84: 621–669; 2012.

[37] Jouguet, P., Kunz-Jacques, S., Leverrier, A., Grangier, P. and Diamanti, E.; *Experimental demonstration of long-distance continuous-variable quantum key distribution*; Nature Photonics 7(5): 378-381; 2013.